

AUV-Based Underwater 3-D Point Cloud Generation Using Acoustic Lens-Based Multibeam Sonar

Hyeonwoo Cho, *Member, IEEE*, Byeongjin Kim, *Member, IEEE*, and Son-Cheol Yu, *Member, IEEE*

Abstract—This study proposes a 3-D seafloor scanning method using sonar images obtained by an acoustic lens-based multibeam sonar (ALMS). ALMS is a forward-looking sonar and is generally used to implement real-time acoustic vision for autonomous underwater vehicles (AUVs). ALMS is originally not a 3-D sonar, and elevation information is eliminated when the 3-D world is mapped to the 2-D sonar image. The proposed method provides a unique analysis of sonar image geometry for extracting the missing elevation information. Different from conventional 3-D reconstruction methods, the proposed method can be continuously executed regardless of the existence of any remarkable features in the sonar image, which means that the proposed method can scan the seafloor during a full-time AUV operation. Once the 3-D seafloor data are obtained, the height of the seafloor can be used for the detection of suspected objects. Therefore, the proposed method enables long-range seafloor object detection, triggering succeeding processes involving the AUV approaching the detected object, and the start of a detailed investigation at a close range.

Index Terms—Acoustic lens-based multibeam sonar (ALMS), seafloor investigation, three-dimensional (3-D) line scanning, 3-D point cloud generation.

I. INTRODUCTION

SEAFLOOR investigation is one of the major tasks of autonomous underwater vehicles (AUVs) [1]–[4]. The hovering-type AUV can be used for the detailed investigation of a specific region of interest [5]–[7]. With the high positioning accuracy of the hovering-type AUV, high-resolution seafloor image maps can be generated using optical imaging [8]–[11].

For the optical investigation, an AUV must maintain a close distance from the seafloor because the visibility from the cameras is limited to short distances [12]. To increase the speed of investigation, the redundant investigation area can be eliminated because the interesting areas may occupy only small regions of the entire seafloor.

An investigation scenario using both acoustic and optical vision systems can be assumed for enhancing work efficiency [13], [14]. Using an acoustic imaging device, the AUV can scan the seafloor from long distances. If a suspected object

Manuscript received January 5, 2017; revised May 11, 2017 and July 28, 2017; accepted September 5, 2017. Date of publication October 5, 2017; date of current version October 11, 2018. This work was supported by the National Research Foundation of Korea Grant funded by the Korea Government (MSIT) (2017R1A5A1014883). (Corresponding author: Son-Cheol Yu.)

Associate Editor: A. Trucco.

The authors are with the Department of Creative IT Engineering, Pohang University of Science and Technology, Pohang 37673, South Korea (e-mail: lighto@postech.ac.kr; kbj0607@postech.ac.kr; sncyu@postech.ac.kr).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JOE.2017.2751139

is detected, then the AUV approaches the object and tries to perform detailed investigation using the optical camera [15], [16]. This scenario eliminates the usage of time for investigating uninteresting areas.

To implement this scenario, sonar-based object detection and a subsequent optical investigation must be executed sequentially, because relocating the AUV to the same point where a suspected object appears is difficult to achieve underwater [16]. Therefore, sonar-based object detection can be executed in real time, and the result can be directly fed to the AUV controller.

The acoustic lens-based multibeam sonar (ALMS) is a high-frequency acoustic video camera that provides live sonar images of underwater scenes [17]. It is suitable for the long-range vision system of hovering-type AUVs. The general purpose of ALMS is real-time monitoring of the underwater scene. Furthermore, the sonar images have been applied for various AUV applications such as underwater navigation, object tracking, pose estimation, and data fusion with optical images [18]–[21].

This study proposes a method for real-time 3-D scanning using ALMS. Most suspected objects protrude from the seafloor, so the elevation of an object can be used for detecting their presence. However, the sonar image produced by the ALMS is the result of mapping a 3-D world to a 2-D pixel array, and this mapping eliminates elevation information [22], [23]. For suspected object detection, the proposed method reconstructs the missing elevation information. This approach can be a more reliable method than 2-D pixel-based detection, which suffers from ambiguity caused by irregular acoustic conditions of the seafloor owing to the motion of the AUV [24].

For 3-D reconstruction, stereo imaging methods for sonar images have been studied [25]–[28]. These methods are based on optical imaging technology [29]. The 3-D shape of an object can be reconstructed based on the pixel correspondence between two sonar images obtained at different locations, and the sonar image model [30], [31]. Several methods use the silhouettes of the shadows in sonar images [32]–[34]. The sonar image region can be divided into three regions: shadows, highlights, and background. Among these regions, the shadow region is a dark region in which no acoustic beam reflection is measured; it is caused by objects blocking the acoustic beam propagation path. Therefore, the lengths of shadows can serve as data for reconstructing the elevation information [34], [35]. These methods assume that remarkable objects exist in the sonar image for matching pixels and segmenting the shadow region. However, for the investigation scenario, 3-D reconstruction must be continuously executed regardless of object existence in the sonar image.

Several studies for 3-D reconstruction in unstructured environments use a special setting of ALMS [36]–[38]. The acoustic beam concentrator is equipped on the sonar, and the sonar is aimed toward the seafloor. This configuration allows for the scanning of a line-shaped region of the seafloor regardless of the object existence. However, this special configuration significantly reduces the field of view of sonar images. Therefore, this method sacrifices the general-purpose use of sonar, namely real-time monitoring of the forward underwater scene.

The proposed method provides 3-D line scan data based on the unique analysis of the sonar geometry in ALMS. The acoustic beams emitted by ALMS have a finite vertical spreading angle [36], [39]. Therefore, the region in which the acoustic beams are reflected is bounded [30]. In sonar images, the bounded region appears as a highlighted region of the seafloor, but the other region (outside the boundary) appears as a shadow region. In this study, we call the highlighted region the effective region, and we call the shadow region the ineffective region.

If there is no protruding object on the seafloor, the boundaries between the effective and ineffective regions for all acoustic beams are located at the equivalent range. However, if a protruding object exists, the highlight of the object extends across the boundary because the acoustic beam reflected by the surface of the object travels a shorter distance than the distance to the boundary. The amount of highlight extension from the effective region into the ineffective region can be used to calculate the missing elevation information.

To generate the full 3-D seafloor, the proposed method needs to utilize the mobility of the AUV. If the global coordinates where the 3-D line scan data are obtained are measured by the AUV positioning system, each 3-D data point from successive sonar images can be assembled in the global coordinates. It forms the 3-D point clouds of the seafloor. The proposed method is developed to perform object detection for AUV investigation, and it requires both AUV position and mobility. Therefore, the proposed method is a collaborative approach suitable for a real-time AUV operation.

II. MOTIVATION

A. Interactive Scenario

Seafloor investigation can be achieved by using optical cameras equipped with a hovering-type AUV [8]–[11]. These optical image-based technologies are used to derive high-resolution seafloor maps based on image mosaicking. Due to limited light visibility in water, the AUV must be maintained at a close distance to the seafloor, which reduces the speed of investigation. Even if strong light enables longer range scanning, the longer range sacrifices the resolution. In addition, maintaining strong light consumes more power, which is a drawback of AUV operation because the AUV has no tether cable for supplying power.

Generally, the hovering-type AUV follows a zigzag trajectory for seafloor investigation, while maintaining a close distance to the seafloor. This guarantees the use of optical imaging devices as shown in Fig. 1(a) [6]. The conventional strategy is still suitable for the investigation and mapping of the entire seafloor.

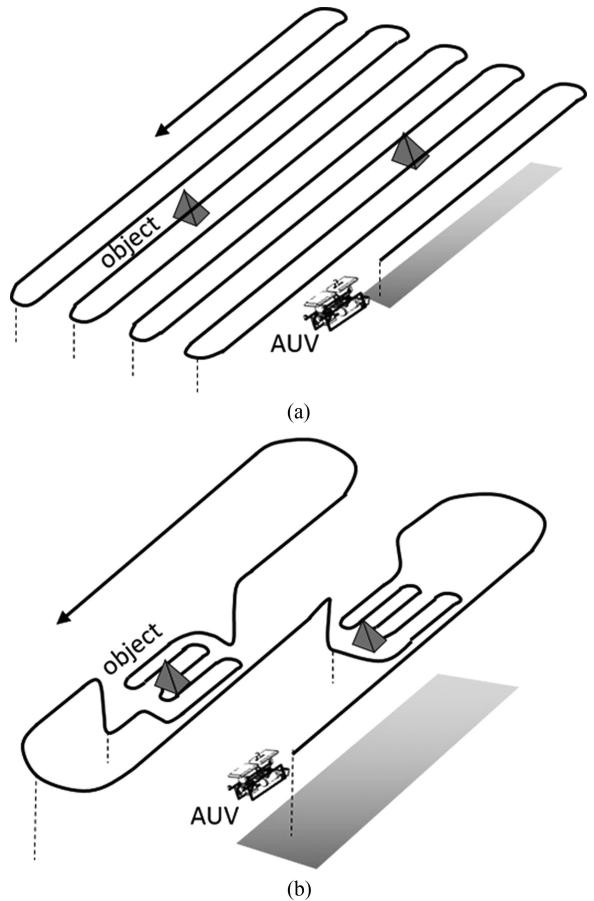


Fig. 1. Underwater investigation scenarios.

However, if the mission objective is an object search and selective investigation, an interactive scenario in which the AUV uses an acoustic camera for long-range scanning can be more effective, as shown in Fig. 1(b). Using sonar images, the AUV roughly scans the seafloor at a long distance away from seafloor [13]–[16]. If a suspected object appears in the sonar image, the AUV approaches close to the seafloor and executes an optical investigation on the area, including the object. After investigation, the AUV returns to its original altitude and continues sonar-based searching.

B. Previous Studies

To implement the interactive scenario, the AUV should automatically detect suspected objects visible in the sonar image. Conventional methods have attempted to find an object in the sonar image based on image processing algorithms. However, image processing suffers from reliability problems because the acoustic characteristics of seafloors are not consistent, owing to the motion of AUV. Therefore, the success rate of detection depends on the environmental conditions [24].

Instead of enhancing the object detection algorithm, this study focuses on reconstructing the elevation of the seafloor using sonar images. Because many suspected objects protrude from the seafloor, object detection becomes a simple task of detecting exceptionally high elevations in the seafloor.

In stereo imaging theory, multiple images obtained at different locations have been used to extract 3-D information from 2-D images [29]. Similarly, 3-D reconstruction can be performed using multiple sonar images [25]–[27]. Furthermore, additional 3-D information such as AUV motion and pose can be extracted from multiple sonar images [19], [20], [28]. Stereo imaging requires feature detection and matching between sonar images, but it may fail due to the low resolution and low signal-to-noise ratio of sonar images. To enhance the degree of feature matching repeatability and distinctiveness, complex feature detection and matching processes are required [30], [31].

Other studies focus on the silhouettes of the shadows in sonar images and adopt the shape from shading (SFS) method [34]. The shadow region of a sonar image is caused by an object that blocks the acoustic beam propagation path. Therefore, the shape of the shadow region can be used for determining the size and edge boundaries of objects in sonar images [34], [35]. The SFS method was originally developed to generate the 3-D information of an object in the field of optical image processing using a single camera [40]. The method was adopted for the 3-D reconstruction of images obtained by sidescan sonar [32], [33]. In recent years, sonar images obtained by ALMS have also been reconstructed by the SFS method [34], [41]. Different from stereo imaging, the SFS method can perform 3-D reconstruction using a single sonar image [34]. However, this method makes several assumptions: the range to the object surface varies monotonically, and the distinct acoustic reflection characteristics of different materials are known in advance [23].

C. Concept of the Proposed Method

Although it is possible to extract 3-D information using existing methods, they are most effective when it is assumed that the target object or area is locked on the sonar image. On the other hand, this study focuses on the interactive scenario that requires the object detection before lock on. The proposed method performs the 3-D seafloor scanning during AUV operation. Through the scanning result, the appearance of object can be detected. By triggering the detection signal, the successive lock-on process and 3-D object investigation are executed even based on the existing methods.

The proposed method first generates 3-D line scan data of the seafloor, which are accumulated together so that a 3-D point cloud of the seafloor can be generated. There has been an analogy of this type of scanning that uses a line laser that is mechanically aligned with an optical camera [42]. The camera captures the seafloor image of the laser line, and an image processing algorithm extracts the line segment from the image. Depending on the distance between the camera and the seafloor, the position of the laser line in the image is changed. Using triangulation, the 3-D line scan data of the seafloor can be calculated with high accuracy [43]. This laser scanning method can be applied to various underwater applications, such as 3-D seafloor reconstruction, AUV navigation, and underwater investigations [42], [44]–[46].

The proposed method is the acoustic image version of 3-D line data generation. Based on the geometrical relationship

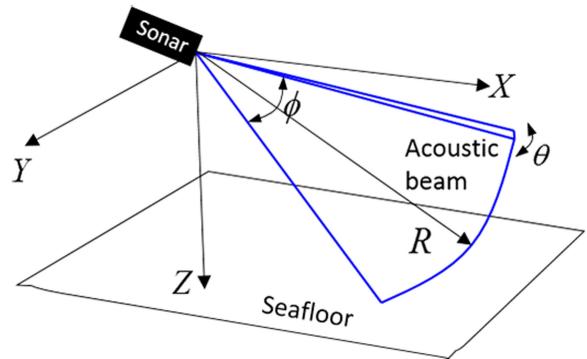


Fig. 2. Shape of the acoustic beam in forward-looking sonar.

between sonar and external parameters, including tilt angle and altitude of ALMS, the elevation of the seafloor can be reconstructed. This means that the 3-D line scan data are obtained per each sonar image. The range of ALMS in underwater is tens of meters, which is much longer than that of an optical camera [39]. Therefore, the proposed method can detect an object protruding from seafloor at father distances than laser line scanning.

Actually, a 3-D seafloor map can be achieved even with multibeam sonar [47] and a special configuration of ALMS [36]–[38]. In particular, multibeam sonar is special equipment used in bathymetry, and allows for large seafloor areas to be scanned. The 3-D seafloor map cannot be used in real time, but generally used as *a priori* knowledge for a human operator before AUV survey [48]. Underwater, it is difficult to re-approach a location where an object was found, after some time has passed [16]. Therefore, if the detection, access, and investigation of objects can be performed continuously, the efficiency of the investigation can be greatly improved. The proposed method can provide 3-D line scan data while obtaining real-time sonar images. This means that the proposed method can be executed simultaneously with existing sonar image-based AUV navigation algorithms [18]–[21]. In addition, the proposed method can be used for object detection in the interactive scenario.

III. HIGHLIGHT EXTENSION

As part of the interactive investigation scenario, the main purpose of the proposed method is to generate 3-D line scan data from a sonar image of ALMS. In this section, we define the effective region and highlight extension of a sonar image under the assumption that the ALMS moves parallel with the seafloor. Based on the definitions, a geometrical model for 3-D line scan generation is proposed.

A. Acoustic Beam Configuration of FLS

ALMS uses fan-shaped acoustic beams and the sonar emits multiple beams in its forward direction. One acoustic beam in the spherical coordinate system is shown in Fig. 2. The beam has a narrow spreading angle along the azimuth direction θ , but it has a wide spreading angle along the elevation angle direction ϕ [49]. The beams are closely aligned with each other along the

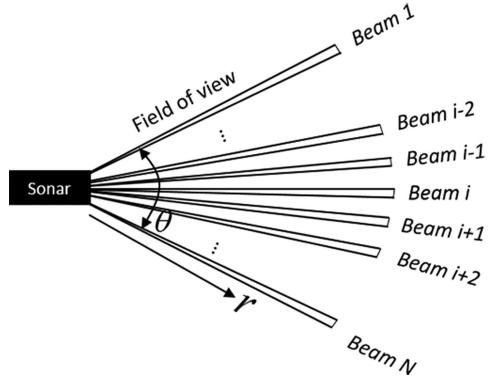


Fig. 3. Acoustic beam configuration in forward-looking sonar.

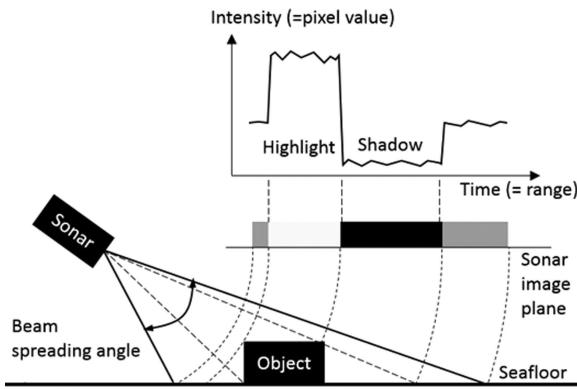


Fig. 4. Acoustic beam reflections over time.

azimuth direction and span the field of view [16], as shown in Fig. 3. In this study, the number of beams is denoted as N .

Once a beam is emitted, it is reflected by both the seafloor and the surface of the object. ALMS measures the reflection intensities over time. The time-of-flight values of each measured reflection depend on the distances to the reflective points. As shown in Fig. 4, the reflections from the seafloor region between the sonar and the object are the first to be measured. After a short time, reflections from the surface of the object are measured. In general, the reflection intensity from the surface of the object is stronger than that of other reflections. Therefore, the object generates a highlight region in the sonar image. After strong reflections, no reflections are measured because the emitted acoustic beam is blocked by the object. This creates the shadow region in the sonar image. The last measurements correspond to the reflection from the seafloor beyond the object.

The time-of-flight values for each reflection can be converted into range values by multiplying with the speed of sound. Moreover, the intensities of each reflection can be converted into gray-scale pixel values [16]. If the number of sampled reflections is assumed to be M , a pixel array of length M is created. Because the number of acoustic beams is assumed to be N , the ALMS obtains N pixel arrays for each beam. Therefore, an $M \times N$ sonar image can be generated by adjoining the pixel arrays. Fig. 5(a) shows the sonar image $S(i, j)$ whose dimensions are $M \times N$ for $1 \leq i \leq M$ and $1 \leq j \leq N$. To explicitly show the acoustic beam configuration, the sonar image is expressed in a

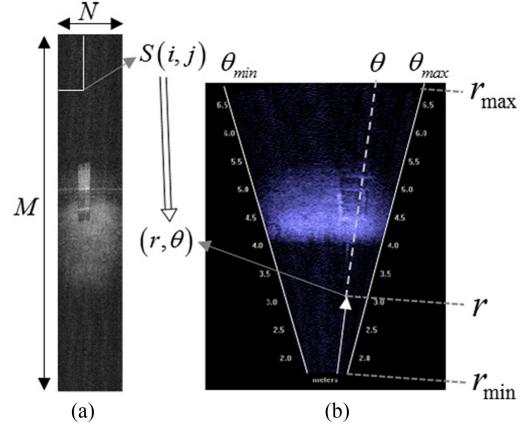


Fig. 5. Sonar image representations. (a) Raw sonar image, (b) Polar coordinate form.

polar-coordinate form, rather than as a rectangular 2-D array, as shown in Fig. 5(b).

According to sonar geometry, a polar coordinate (r, θ) can be assigned to each pixel index (i, j) . As shown in Fig. 5(b), the field-of-view of the sonar image can be defined as θ_{\min} and θ_{\max} . In addition, the minimum and maximum ranges of the sonar image can be defined as r_{\min} and r_{\max} , respectively. In this situation, a pixel index (i, j) can be converted into (r, θ) using the following equations [50]:

$$r = r_{\min} + \epsilon_r (i - 0.5) \quad (1)$$

and

$$\theta = \theta_{\min} + \epsilon_\theta (j - 0.5) \quad (2)$$

where ϵ_r and ϵ_θ are the radial and angular resolutions, respectively, and are defined as

$$\epsilon_r = \frac{r_{\max} - r_{\min}}{M} \quad (3)$$

and

$$\epsilon_\theta = \frac{\theta_{\max} - \theta_{\min}}{N}. \quad (4)$$

Equations (1) and (2) are based on the ideal acoustic beam configuration. In practice, the error of the beam configuration can be calibrated [51].

B. Loss of Elevation Angle

The spherical coordinate system, which is defined as a triplet (r, ϕ, θ) , can be used for describing positions in the 3-D world. The elevation angle ϕ cannot be obtained from the pixel index (i, j) because the sonar image is a mapping of 3-D spherical coordinates to 2-D polar coordinates, which are defined by (r, θ) .

In Fig. 6, there are three positions that are located at the same range but different elevation angles. Based on sonar geometry, the three points are mapped to the same pixel in the sonar image. Therefore, the elevation angle ϕ cannot be calculated from pixel indices of the sonar image [18], [34]. In Cartesian coordinates, this situation is equivalent to the height of the object being unknown.

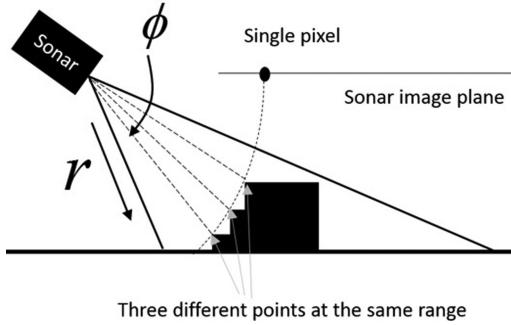


Fig. 6. Ambiguity of elevation angle.

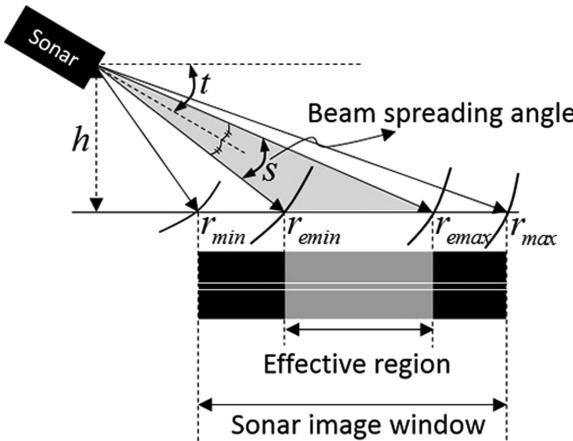


Fig. 7. Effective region and sonar image window.

C. Effective Region

In this section, we define the effective region of the sonar image and also the highlight extension, which forms the basis of the proposed 3-D scanning. The acoustic beams of ALMS have a finite vertical beam spreading angle s , as shown in Fig. 7. If the altitude h and tilt angle t of the ALMS are also given, the region in which the acoustic beam is actually reflected is bounded by the ranges r_{emin} and r_{emax} , which are calculated as

$$r_{\text{emin}} = \frac{h}{\sin(t + 0.5s)} \quad (5)$$

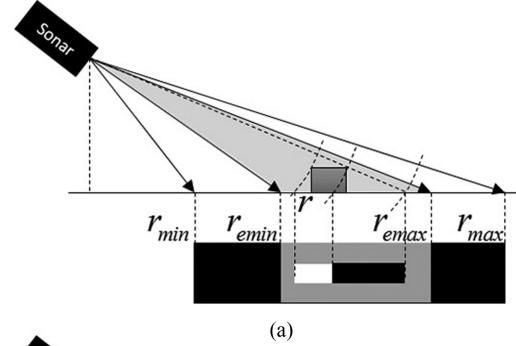
and

$$r_{\text{emax}} = \frac{h}{\sin(t - 0.5s)}. \quad (6)$$

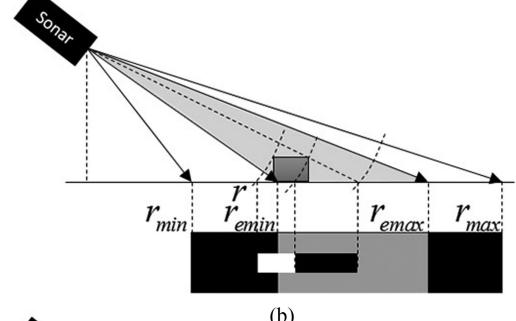
In this study, the region between these limits is called the effective region. In general, the effective region has brighter pixels than the shadow region because of the acoustic beam reflection from the seafloor.

Independent of r_{emin} and r_{emax} , the minimum and maximum ranges of the sonar image, which are r_{min} and r_{max} , respectively, can be specified by a human operator. In this study, the region bounded by r_{min} and r_{max} is called the sonar image window, as shown in Fig. 7.

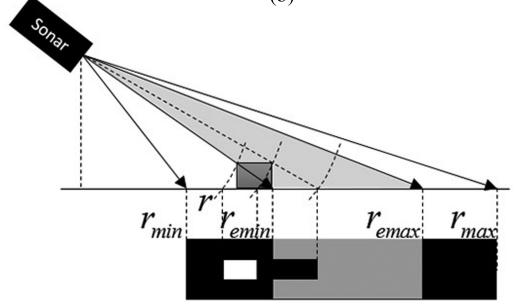
If the sonar image window is wider than the effective region, only the middle part of the sonar image has high pixel values that coincide with the highlights of the effective region, as shown



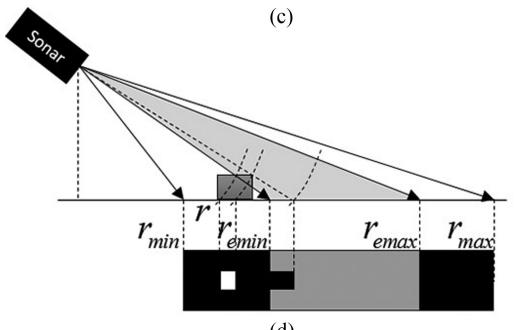
(a)



(b)



(c)



(d)

Fig. 8. Highlight extension caused by the relative distance change between forward-looking sonar and object. (a) Before highlight extension, (b) Beginning of highlight extension, (c) Critical point, (d) After critical position.

in Fig. 7. The other regions of the sonar image have low pixel values because no acoustic beam reflection exists in the other regions. In this study, the other regions are called the ineffective region.

D. Highlight Extension

Assume that an object is located on the seafloor and that the ALMS device approaches the object. Fig. 8(a) shows a situation in which the object is just beginning to appear in the sonar

image. In this image, the entire highlight region of the object is within the effective region.

Fig. 8(b) shows a situation in which the ALMS approaches the object at a closer distance. In this situation, the highlight region of the object extends beyond the effective region. The surface of the object at which the acoustic reflection occurs is closer than r_{emin} ; therefore, the highlight of the object is mapped to a closer range than the boundary between the effective and ineffective regions. In this study, this situation is called highlight extension.

When the ALMS device moves further, the amount of the highlight extension increases, and eventually, the front part of the object meets one of the edges of the spreading acoustic beam of the FLS device, as shown in Fig. 8(c). In this study, this location is called the critical point.

Fig. 8(d) shows a situation in which the front of the surface of the object passes the edge of the spreading acoustic beam of the ALMS device after the critical point. The amount of highlight extension is the same as the amount of highlight extension at the critical point because the front part of the object is outside the range of the spreading acoustic beam.

As long as the ALMS device moves in the forward direction, the ALMS must meet the critical point, and highlight extension will be constant after the critical point. The amount of highlight extension can be converted into the height of the object.

IV. THREE-DIMENSIONAL POINT CLOUD GENERATION

In this section, we present a technique for detecting highlight extension and describe the mathematical equations used to calculate the height of the object based on the length of the extended highlight.

A. Detection of Highlight Extension

As mentioned above, the ineffective region has low pixel values and appears as a shadow region. On the other hand, the extended highlight has high pixel values. Because of their high contrast, the highlight extension causes extreme pixel value increases in the ineffective region. Therefore, highlight extension can be readily detected by applying difference filters, which yield a peak at the point at which extreme pixel value variation occurs.

When the number of row indices of the sonar image is M , r_{\min} and r_{\max} correspond to the row pixel indices 1 and M , respectively. Therefore, the row indices corresponding to r_{emin} and r_{emax} , which are denoted as M_{emin} and M_{emax} , respectively, can be calculated as

$$M_{\text{emin}} = M \left(\frac{r_{\text{emin}} - r_{\min}}{r_{\max} - r_{\min}} \right) \quad (7)$$

and

$$M_{\text{emax}} = M \left(\frac{r_{\text{emax}} - r_{\min}}{r_{\max} - r_{\min}} \right). \quad (8)$$

When M_{emin} and M_{emax} are not integers, a rounding operation can be applied to them.

Based on M_{emin} and M_{emax} , the sonar image $S(i, j)$ can be divided into three subparts, i.e., $S_1(i, j)$, $S_2(i, j)$, and $S_3(i, j)$

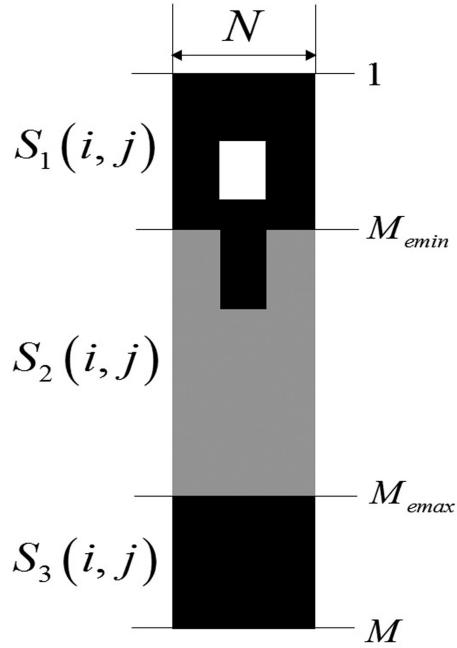


Fig. 9. Sonar image separation.

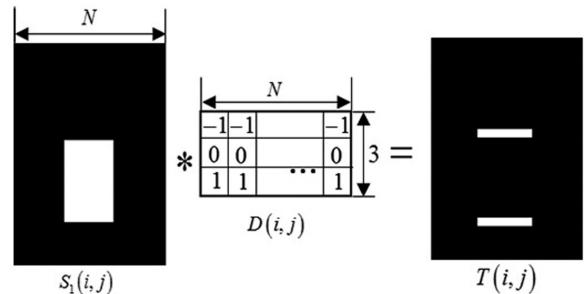


Fig. 10. Difference filter for detecting the extended highlight.

whose row index ranges satisfy $1 \leq i < M_{\text{emin}}$, $M_{\text{emin}} \leq i \leq M_{\text{emax}}$, and $M_{\text{emax}} < i \leq M$, respectively, as shown in Fig. 9. Because the highlight extension occurs in $S_1(i, j)$, the difference filter is applied only to $S_1(i, j)$ rather than the entire sonar image.

As shown in Fig. 10, the difference filter is defined using a $3 \times N$ kernel $D(i, j)$ whose values are set to -1 , 0 , and 1 for $i = -1$, $i = 0$, and $i = 1$, respectively, for all j . Using the correlation operation between $S_1(i, j)$ and $D(i, j)$, the filtered image, $T(i, j)$, can be derived as

$$T(i, j) = \sum_{k=-1}^1 S_1(i+k, j)D(k, j). \quad (9)$$

Note that the correlation is a 1-D operation that shifts only the row index, because each column of the sonar image is treated as a separate pixel array (based on the sonar image generation mechanism); the detection of the highlight extension can be independently executed for each column.

Ideally, one of the columns of $T(i, j)$ that includes the extended highlight has one positive peak, which corresponds to

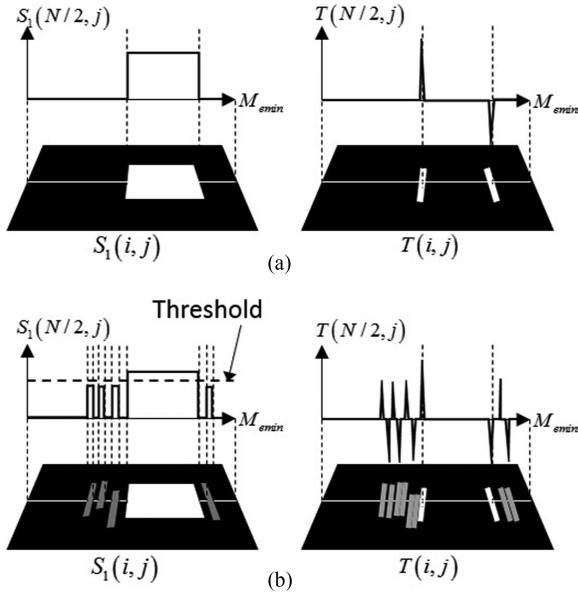


Fig. 11. Extended highlight front-end determination. (a) Ideal case. (b) Practical case.

the rising edge of the extended highlight of $S_1(i,j)$, as shown in Fig. 11(a). A negative peak can also appear, and it corresponds to the falling edge of the extended highlight. Therefore, the index of the positive peak can be selected as the front end of the highlight extension.

In practical cases, the filtered image $T(i,j)$ may have many peaks that are not caused by the extended highlight but by undesired noise, as shown in Fig. 11(b). Applying a smoothing filter is one solution for mitigating noise effects. However, there is another practical problem, namely detecting the front end of the extended highlight. The peak derived by the difference filter is due to the relative pixel value difference of adjacent pixels; therefore, many peaks may occur, even if pixel values are not high enough to be considered part of the extended highlight.

To solve this problem, the highlight pixels in $S_2(i,j)$ values can be used as threshold values for determining whether the pixel of the detected front end is the highlight. At first, the pixel of the highest positive peak is selected as the candidate. If the pixel value is equal to or higher than the threshold value, the detected peak is determined to be the front end of the extended highlight. Otherwise, the detected peak is ignored and the situation is treated as having no highlight extension.

B. Elevation Angle Calculation

The highlight detection process is executed for all N columns of $T(i,j)$. As a result, we can obtain N indices for each front end of the highlight extension for each column. We denote the array of indices as $m(j)$ for $1 \leq j \leq N$. If highlight extension is detected in the j th column, $m(j)$ is set to the index of the front end of the highlight. If no peak is detected in the j th column, $m(j)$ is set to M_{emin} , which implies that the j th column does not include an object higher than the seafloor.

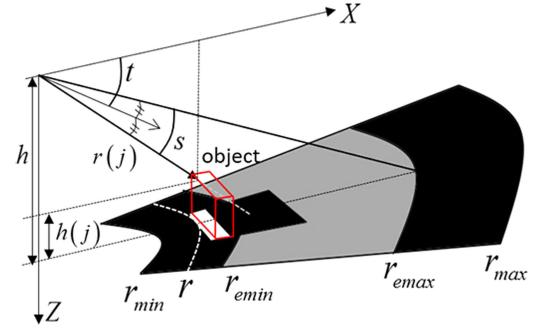


Fig. 12. Geometry used for height calculation.

The range $r(j)$ corresponding to $m(j)$ can be calculated as

$$r(j) = r_{min} + (r_{max} - r_{min}) \frac{m(j)}{M}. \quad (10)$$

Finally, the height of an object that reflects the j th acoustic beam, which is denoted as $h(j)$, can be derived as

$$h(j) = h - r(j) \sin(t + 0.5s) \quad (11)$$

for $1 \leq j \leq N$, as shown in Fig. 12. Therefore, the proposed method calculates the heights corresponding to the sonar image pixels whose indices are $S(m(j), j)$ for all j . In the spherical coordinate system, the elevation angle $\phi(j)$ can be calculated as

$$\phi(j) = \sin^{-1} \left(\frac{h - h(j)}{r(j)} \right). \quad (12)$$

Based on (2), the azimuth angle of the pixels of $S(m(j), j)$ can be calculated as

$$\theta(j) = \theta_{min} + \epsilon_\theta (j - 0.5). \quad (13)$$

As a result, the 3-D coordinate corresponding to the pixel $S(m(j), j)$ can be expressed as $(r(j), \phi(j), \theta(j))$ in spherical coordinates. This coordinate can also be converted into Cartesian coordinates $(x(j), y(j), z(j))$ using coordinate conversion equations

$$x(j) = r(j) \times \cos \phi(j) \times \cos \theta(j) \quad (14)$$

$$y(j) = r(j) \times \cos \phi(j) \times \sin \theta(j) \quad (15)$$

$$z(j) = r(j) \times \sin \phi(j). \quad (16)$$

C. Three-Dimensional Point Cloud Generation

One pixel in each column of a sonar image is converted to 3-D coordinates by the proposed method. Therefore, the number of 3-D coordinates is N and the coordinates constitute line scan data of the seafloor. If an ALMS-equipped AUV successively obtains the line scan data from each sonar image, the 3-D point cloud of the seafloor can be generated by accumulating line scan data in the same coordinate system, as shown in Fig. 13.

This study defines the position of the AUV in global coordinates, as shown in Fig. 14. The position of the AUV when the k th sonar image is obtained is denoted as $(\bar{X}_k, \bar{Y}_k, \bar{Z}_k)$. In addition, the pan, tilt, and yaw angles of the AUV can be denoted as α_k , β_k , and γ_k , respectively. The coordinates of the 3-D line scan data by the proposed method are expressed as the AUV

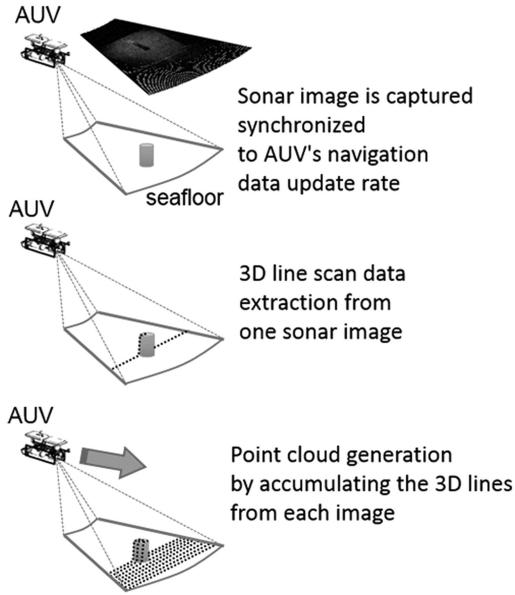


Fig. 13. Three-dimensional point cloud of seafloor generation using AUV.

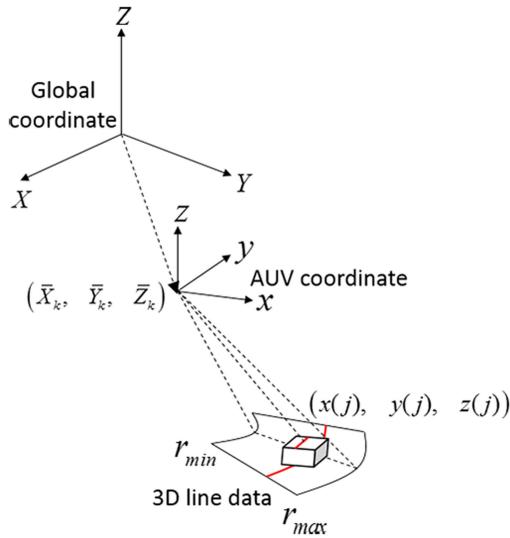


Fig. 14. Three-dimensional scan data in AUV coordinate and its relationship to global coordinates.

coordinates whose origin is (X_k, Y_k, Z_k) in global coordinates, as shown in Fig. 14.

If the coordinates of the 3-D line scan data obtained from the k th sonar image are defined as $x_k(j)$, $y_k(j)$, and $z_k(j)$ for all columns, the coordinates can be converted to global coordinates as

$$\begin{bmatrix} X_k(j) \\ Y_k(j) \\ Z_k(j) \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_k(j) \\ y_k(j) \\ z_k(j) \end{bmatrix} + \begin{bmatrix} \bar{X}_k \\ \bar{Y}_k \\ \bar{Z}_k \end{bmatrix} \quad (17)$$

where $X_k(j)$, $Y_k(j)$, and $Z_k(j)$ are the global coordinates of 3-D line scan data from the k th sonar image. \mathbf{R} is the rotation

matrix defined as $\mathbf{R} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z$. Here

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_k & -\sin \alpha_k \\ 0 & \sin \alpha_k & \cos \alpha_k \end{bmatrix} \quad (18)$$

$$\mathbf{R}_y = \begin{bmatrix} \cos \beta_k & 0 & \sin \beta_k \\ 0 & 1 & 0 \\ -\sin \beta_k & 0 & \cos \beta_k \end{bmatrix} \quad (19)$$

and

$$\mathbf{R}_z = \begin{bmatrix} \cos \gamma_k & -\sin \gamma_k & 0 \\ \sin \gamma_k & \cos \gamma_k & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (20)$$

The 3-D line scan data $X_k(j)$, $Y_k(j)$, and $Z_k(j)$ for all j , synchronized with the sonar image frame rate, can be accumulated in the global coordinates. This accumulation generates the 3-D point cloud data of the seafloor.

V. EXPERIMENTS

Two experiments were conducted to verify the performance of the proposed method. The first experiment was executed in a wave basin to test the accuracy of the height measured using the proposed method. The second was a field experiment to verify the feasibility of the proposed method for use during AUV operation.

In both experiments, dual frequency identification sonar (DIDSON), which is one of the ALMS devices, was used [39]. The beam spreading angle of DIDSON is 14° , and the number of acoustic beams is 96, which means that $s = 14^\circ$ and $N = 96$ [39]. The number of rows is 512; therefore, $M = 512$. The field of view of DIDSON along the azimuth direction is 29° , which means that $\theta_{\min} = -14.5^\circ$ and $\theta_{\max} = 14.5^\circ$ [39].

A. Indoor Experiment

The wave basin is a long water tank that a carriage can move over at a specific constant speed, as shown in Fig. 15(a). For this experiment, we made a rigid frame for mounting the DIDSON and Doppler velocity log (DVL), and they were attached under the carriage, as shown in Fig. 15(a). The DVL was used for measuring the altitude of DIDSON from the bottom of the water tank. The tilt angle of DIDSON was set to 45° downward. The sonar image window limits, r_{\min} and r_{\max} , were set to 1.6760 and 6.6760 m, respectively. The six plastic baskets are placed in the bottom of the wave basin, as shown in Fig. 15(b). The length, width, and height of the basket are 720, 470, and 330 mm, respectively.

In this setting, we moved the carriage at a constant speed of 0.5 m/s. While moving the carriage, successive sonar images of the bottom of the wave basin were obtained using DIDSON. The number of total sonar images obtained in this experiment was 368. The representative sonar images of the plastic baskets captured in this experiment are shown in Fig. 16.

Synchronized with sonar image acquisition, the altitudes from the bottom of the wave basin were also measured by DVL, and

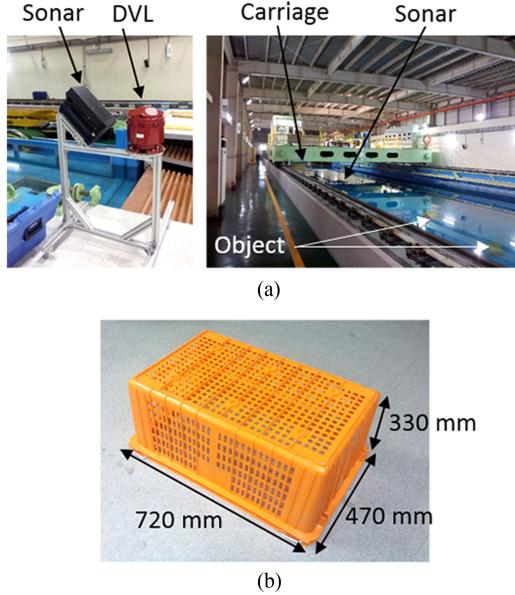


Fig. 15. Setup of experiment in wave basin. (a) Experiment setting. (b) Object size.

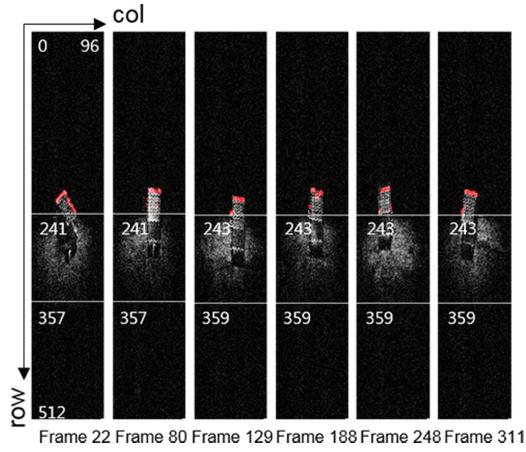


Fig. 16. Representative sonar images of plastic baskets in wave basin.

the measured altitudes are shown in Fig. 17(a). Based on the measurements, the boundary of the effective region, M_{emin} and M_{emax} , were calculated and they are graphed in Fig. 17(b). The variation of M_{emin} and M_{emax} show the same tendency as a variation of measured altitudes because M_{emin} and M_{emax} depend on the altitude.

Successive sonar images obtained while passing over the second object are shown in Fig. 18(a). In the first image, the entire highlight of the object is located in the effective region. In the following images, the object highlight begins to extend into the effective region and vanishes after DIDSON has completely passed the object.

For each sonar image, the highlight extension was detected and the pixel coordinates of the front end of the highlight extension were converted into 3-D line scan data. The heights of the 3-D line scan data of the sonar images are graphed in Fig. 18(b). We can see a significant height increase when the plastic basket

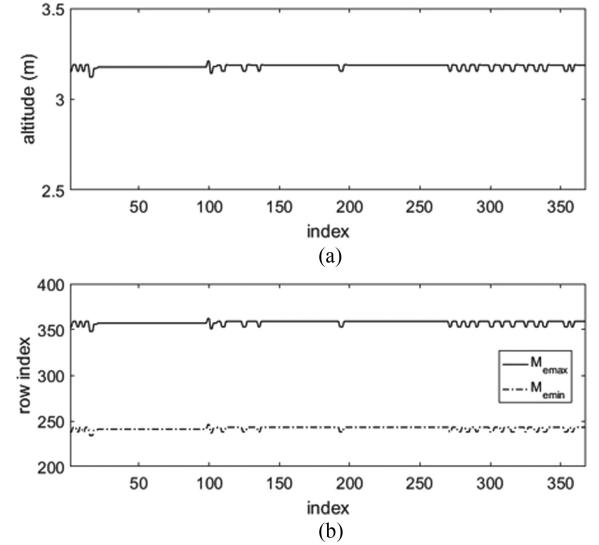


Fig. 17. Measured altitudes and calculated effective region boundary during the experiment.

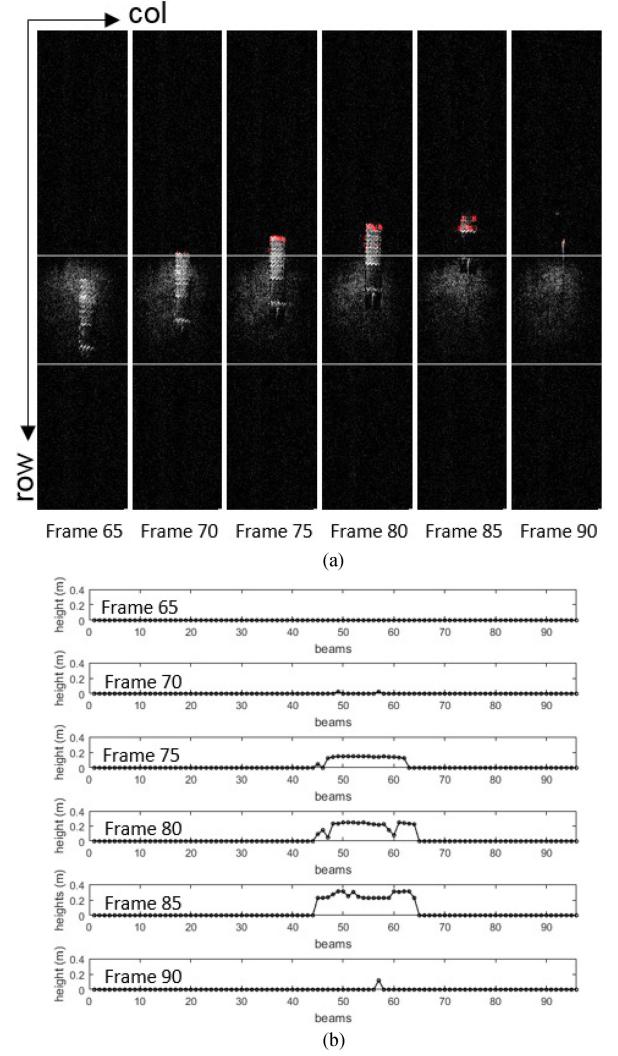


Fig. 18. Three-dimensional line scan data generation of the second object. (a) Successive sonar images of the second object. (b) Three-dimensional line scan data generation of the second object.

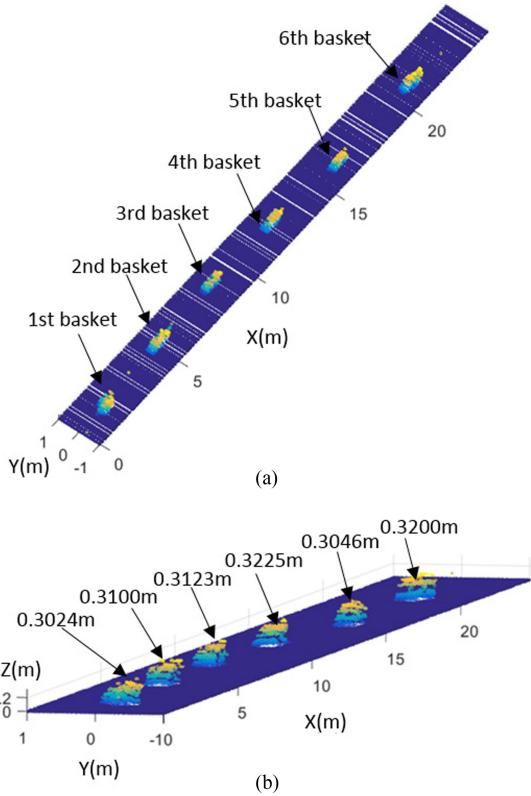


Fig. 19. Three-dimensional point clouds of plastic baskets on the bottom of the wave basin. (a) Perspective view. (b) Height measurement

appeared. This increase can be easily detected and be used as the trigger signal to indicate the appearance of a suspected object.

For 3-D cloud point generation, we define a Cartesian coordinate where the positive X -axis is set to the moving direction of DIDSON. The positive Z direction is set to the perpendicular direction from the bottom of the water tank. The 3-D line scan data of each sonar image are mapped to the Cartesian coordinate, and the 3-D point cloud for the bottom of the wave basin can be generated as shown in Fig. 19. We can clearly see the 3-D cloud points of the plastic baskets. The heights of the plastic baskets were 0.3023, 0.3100, 0.3123, 0.3225, 0.3046, and 0.3200 m. Because the true height is 0.33 m, the errors are $-0.0277, -0.0200, -0.0177, -0.0075, -0.0254$, and -0.0100 m, respectively. The average of the errors is -0.0181 m.

B. Effect of External Measurement Error

The proposed method uses sonar parameters and external measurements for calculating 3-D line scan data. In this section, we analyze the effect of their error on detectability and accuracy. First, detectability means whether the proposed method can calculate the object 3-D line scan data, even if an error exists. Fig. 20 shows the sonar geometry of the proposed method with external parameter errors. As a sonar parameter, s' is the beam spreading angle with error. As external measurements, h' and t' are the altitude and tilt angle with error. H is the true height of the object. Using (5), the erroneous range of effective region boundary r'_{emin} is calculated as $r'_{\text{emin}} = (h'/\sin(T'))$, where

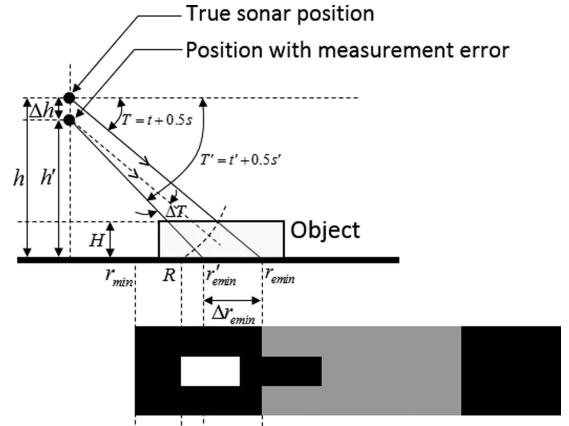


Fig. 20. Proposed sonar geometry with external parameter errors.

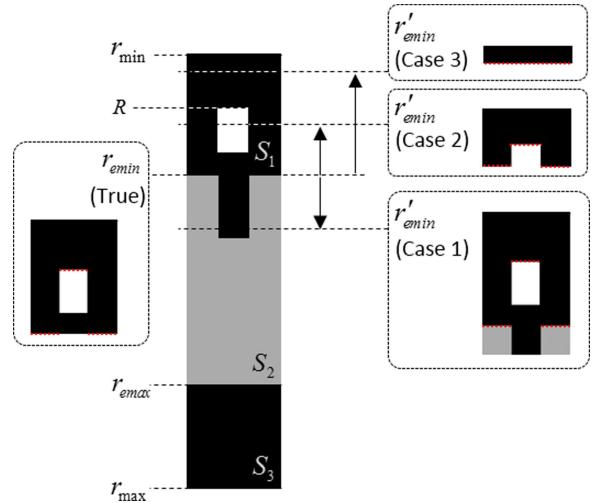


Fig. 21. Comparison of each S_1 bounded by r'_{emin} .

$T' = t' + 0.5s'$. As shown in Fig. 20, r'_{emin} is different from the true r_{emin} , because of the errors.

In the proposed method, the sonar image is divided into S_1 , S_2 , and S_3 , based on the effective region. The highlight extension of the object is detected in S_1 . r'_{emin} can extend or reduce the area of S_1 , as shown in Fig. 21, for cases 1 and 2. In particular, if r'_{emin} is less than R , which is the range to the edge of the object's highlight extension, the highlight extension does not belong to S_1 , as shown in Fig. 21 for case 3. Therefore, the proposed method cannot detect the highlight extension in S_1 . This implies an undetectable situation.

To determine the detectability condition, with respect to the external parameter error, we define the differential of altitude and tilt angle as dh and dT , respectively. Equation (5) is rewritten as $r_{\text{emin}} = (h/\sin(T))$, where $T = t + 0.5s$; therefore, the differential of dr_{emin} can be written as follows:

$$dr_{\text{emin}} = \frac{\partial r_{\text{emin}}}{\partial h} dh + \frac{\partial r_{\text{emin}}}{\partial T} dT \quad (21)$$

$$= \frac{1}{\sin(T)} dh - h \frac{\cos(T)}{\sin^2(T)} dT. \quad (22)$$

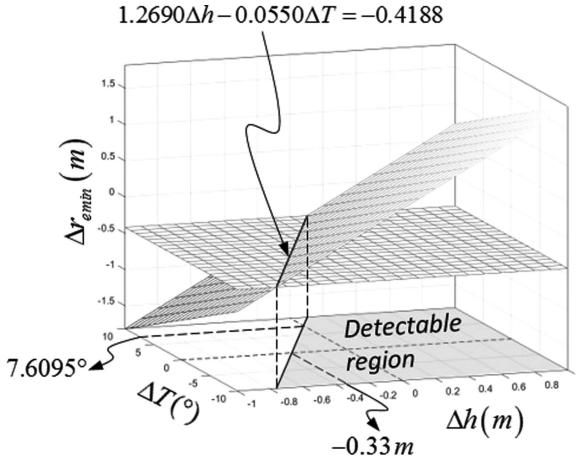


Fig. 22. Detectability with respect to Δh and ΔT .

Using (22), the error of r_{emin} around a given h and T can be approximated as follows:

$$\Delta r_{\text{emin}} \approx \frac{1}{\sin(T)} \Delta h - h \frac{\cos(T)}{\sin^2(T)} \Delta T \quad (23)$$

where $\Delta h = h' - h$, $\Delta T = T' - T$, and $\Delta r_{\text{emin}} = r'_{\text{emin}} - r_{\text{emin}}$. As shown in Fig. 21, S_1 includes the object's highlight extension while $r'_{\text{emin}} > R$. Therefore, the condition of detectability is derived as

$$\frac{1}{\sin(T)} \Delta h - h \frac{\cos(T)}{\sin^2(T)} \Delta T > R - r_{\text{emin}}. \quad (24)$$

For example, if we use the indoor experiment settings, (24) becomes $1.2690\Delta h - 0.0550\Delta T > 3.6173 - 4.0361 = -0.4188$. Under the assumption that Δh and ΔT are varied for $-1.0 \leq \Delta h \leq 1.0$ m and $-10^\circ \leq \Delta T \leq 10^\circ$, respectively, the graph of the detectability condition is shown in Fig. 22. If $\Delta T = 0$, the object is detectable when $\Delta h > -0.33$ m. Additionally, if $\Delta h = 0$, the object is detectable when $\Delta T < 7.6095^\circ$.

To verify these conditions, we generated 3-D point clouds after intentionally adding the altitude and tilt angle errors to the indoor experiment data. Fig. 23 shows the 3-D point clouds when the altitude error Δh was set to -0.43 , -0.33 , and -0.23 m. Here, -0.33 m is the threshold value of Δh for the detectability condition. As expected, the plastic baskets were detected at $\Delta h = -0.23$ m, even if their height was lower than the true height. At $\Delta h = -0.33$ and -0.43 m, the plastic baskets were not detected. This proves that the theoretical detectability condition is $\Delta h > -0.33$ m.

Similarly, Fig. 24 shows the 3-D point clouds when the tilt angle error ΔT was set to 6.6° , 7.6095° , and 8.6° . Here, 7.6095° is the threshold value of ΔT for the detectability condition. According to the results, all six plastic baskets were detected at $\Delta T = 6.6^\circ$, even if their height was lower than the true height. Theoretically, at $\Delta T = 7.6095^\circ$ and 8.6° , all baskets should not be detectable, but some of them were detected. However, the heights of the detected baskets were close to zero, and as ΔT increased, the number of detected baskets decreased.

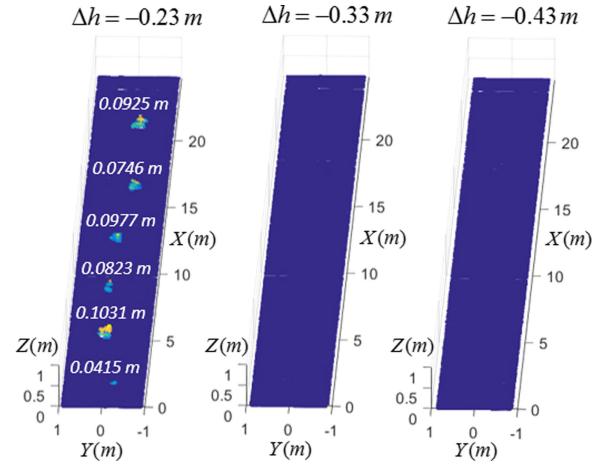


Fig. 23. Three-dimensional point clouds of plastic baskets with altitude errors of -0.23 , -0.33 , and -0.43 m.

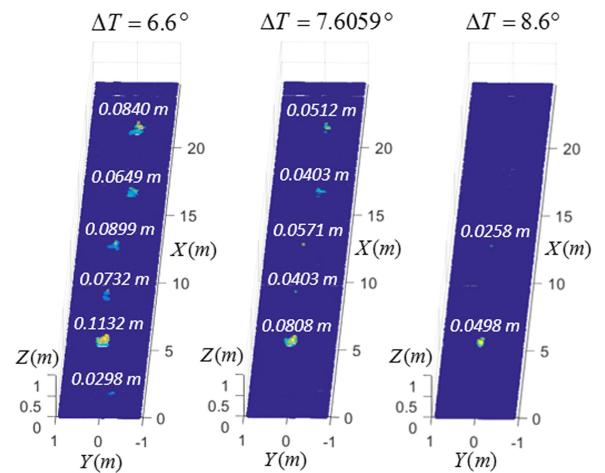


Fig. 24. Three-dimensional point clouds of plastic baskets with tilt angle errors of 6.6 , 7.6095 , and 8.6° .

This result verifies that the theoretical detectability condition of $\Delta T < 7.6095^\circ$ is valid.

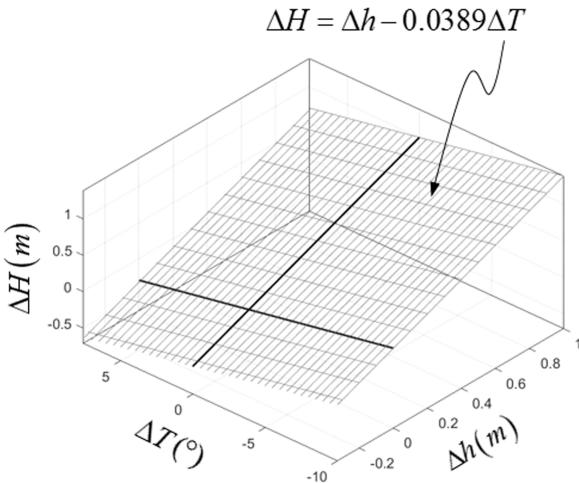
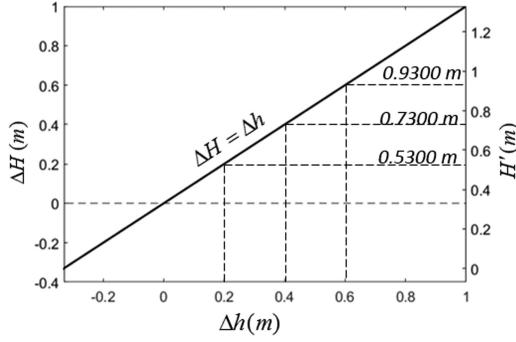
We can analyze the accuracy of the proposed method when the detectability condition is satisfied. According to (11), the erroneous height of object H' , with respect to h' and T' , is given by

$$H' = h' - R \sin(T'). \quad (25)$$

Here, R is obtained from the result of the highlight extension detection algorithm, which is based on image processing; therefore, R is independent of external measurements. Based on (11), the true height of the object H is rewritten as $H = h - R \sin(T)$, and the differential of H , i.e., dH , with respect to dh and dT , is given by

$$dH = \frac{\partial H}{\partial h} dh + \frac{\partial H}{\partial t} dt \quad (26)$$

$$= dh - R \cos(T) dT. \quad (27)$$

Fig. 25. Height error with respect to Δh and ΔT .Fig. 26. Height error and erroneous height with respect to Δh .

Using (27), the error of H , around a given h and T , can be approximated as follows:

$$\Delta H \approx \Delta h - R \cos(T) \Delta T \quad (28)$$

where $\Delta H = H' - H$. The amount of height error caused by Δh and ΔT is independent of the object height; therefore, the larger the actual object height, the smaller the percentage error for the same ΔH . Additionally, we can know that the altitude error Δh is directly transferred to height error ΔH . However, the tilt angle error ΔT is multiplied by $R \cos(T)$, and added to the true height of object H . Therefore, a large T can reduce the effect of ΔT ; however, this makes ALMS look almost downward. Therefore, this causes a reduction in the object detection range.

Based on the indoor experiment settings, (28) is given by $\Delta H \approx \Delta h - 0.0389\Delta T$. While $-0.33 \leq \Delta h \leq 1.0$ m and $-10^\circ \leq \Delta T \leq 7.6095^\circ$, the height error ΔH is derived as shown in Fig. 25.

Fig. 26 shows the graph of ΔH versus Δh , when ΔT is assumed to be zero. As mentioned, the height error ΔH is the same as the altitude error Δh . When the true height H is 0.33 m, the erroneous height to be calculated by the proposed method, with respect to the altitude error, is indicated on the right vertical axis of Fig. 26.

Fig. 27 shows the graph of ΔH versus ΔT , when Δh is assumed to be zero. The height error ΔH is the tilt angle error

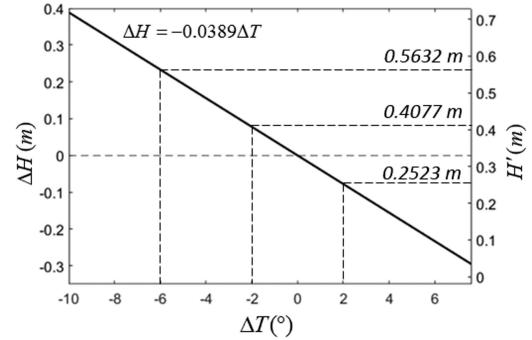
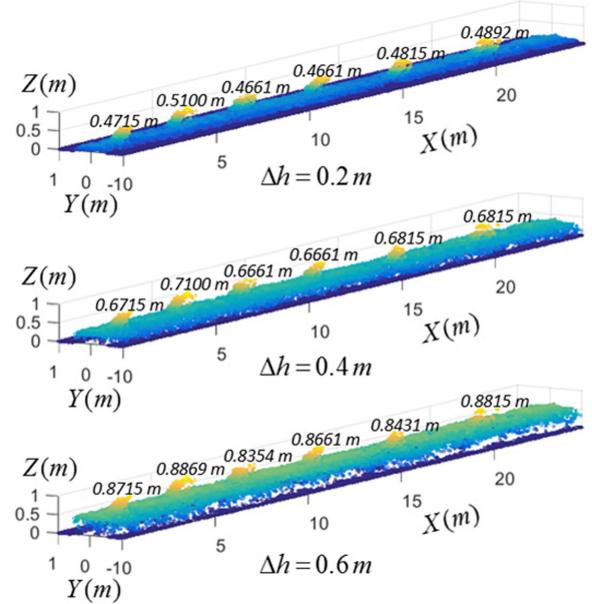
Fig. 27. Height error and erroneous height with respect to ΔT .

Fig. 28. Comparison of 3-D point clouds of each plastic basket with altitude errors.

multiplied by the slope -0.0389 . When the true height H is 0.33 m, the erroneous height to be calculated by the proposed method with tilt angle error is indicated on the right vertical axis of Fig. 27.

To verify the theoretical results shown in Figs. 26 and 27, 3-D point clouds were generated by adding the altitude errors of 0.2, 0.4, and 0.6 m, and tilt angle errors of -2° , 2° , and 6° to the indoor experiment data. Figs. 28 and 29 show the results with altitude and tilt angle errors, respectively. For each result, the mean value of the derived heights is summarized in Table I. The theoretical height values, with respect to given error cases, are extracted from Figs. 26 and 27, and are also written in Table I. According to Table I, the difference between the experimental and theoretical heights is less than a few centimeters. This confirms that the theoretical results are almost identical to the experimental results.

So far, we discussed the effect of h and T errors. Here, T is the term into which t and s are combined. Generally, we can assume that the error of t is much greater than that of s , because t is a time-varying measurement whereas s is one of the sonar parameters specified by ALMS manufacturers. Moreover, the

TABLE I
COMPARISON BETWEEN EXPERIMENTAL AND THEORETICAL RESULTS WITH EXTERNAL PARAMETER ERROR

	Height (H')	Altitude error (Δh)			Tilt angle error (ΔT)		
		0.2 m	0.4 m	0.6 m	2.0°	-2.0°	-6.0°
Experiment (m)	0.4807	0.6795	0.8641	0.2218	0.3684	0.5338	
	0.5300	0.7300	0.9300	0.2523	0.4077	0.5632	
	-0.0493	-0.0506	-0.0659	-0.0305	-0.0394	-0.0294	

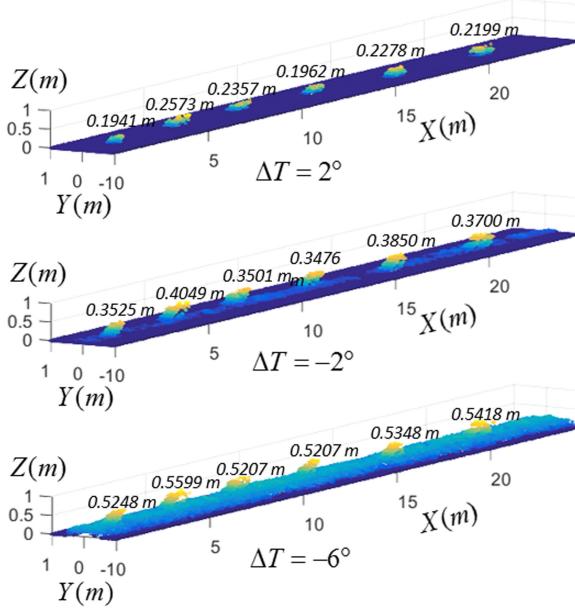


Fig. 29. Comparison of 3-D point clouds of each plastic basket with tilt angle errors.

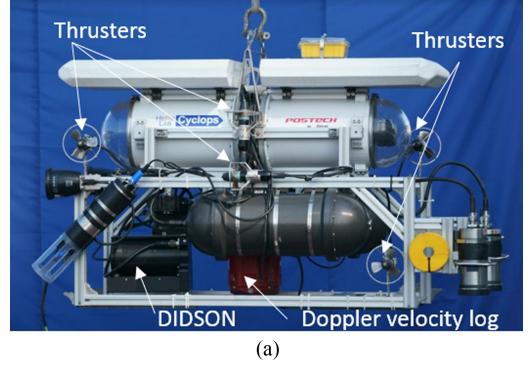
sonar parameters can be calibrated to remove their intrinsic error [51].

There are other sonar parameters, such as r_{\min} and r_{\max} . Parameter r_{\min} does not affect the results; however, it determines the height limit of the object, which the proposed method can compute. Smaller r_{\min} makes the S_1 area wider; therefore, the higher object can also be calculated. Parameter r_{\max} is also considered in the proposed sonar geometry for the purpose of generalization; however, r_{\max} is not actually related to the proposed method, since the proposed method uses only S_1 , which is independent of r_{\max} .

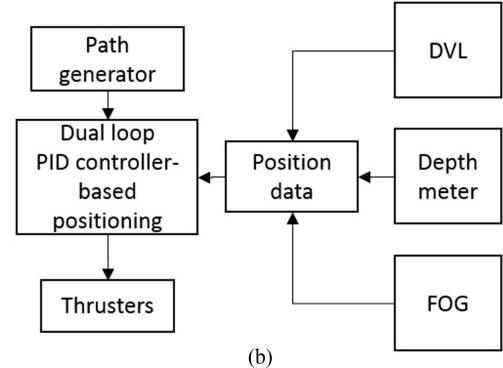
Other external parameters such as pan and yaw angle errors, which are not explicitly considered in this study, also cause the error of the calculated height. In addition, the proposed method is developed under the assumption that the ALMS is parallel to the seafloor. However, if the ALMS is not parallel to the seafloor, which means the seafloor is slanted, it can also cause incorrect height. Therefore, as a future work, specific methods for treating their effect can improve the accuracy of the proposed method.

C. Field Experiment

For a field experiment, a hovering-type AUV was used, as shown in Fig. 30(a). The AUV has eight thrusters: two for surge,



(a)



(b)



(c)

Fig. 30. Hovering-type AUV used for experiment. (a) Hardware configuration. (b) Position control system configuration, AUV image of field experiment.

four for sway, and two for heave motions [6]. As shown in Fig. 30(b), the dual-loop PID controller controls the thrusters to minimize the measured position and desired positions of the path of the AUV [6]. As a result, accurate positioning and station keeping are available under current. The heading of the AUV is measured by a fiber optic gyro. The X coordinate, Y coordinate, and the altitude of the AUV are measured by DVL.

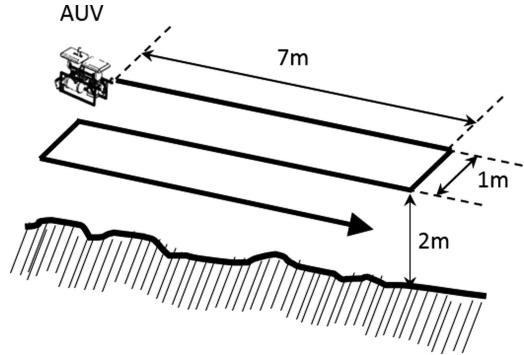


Fig. 31. Path of AUV for field experiment.

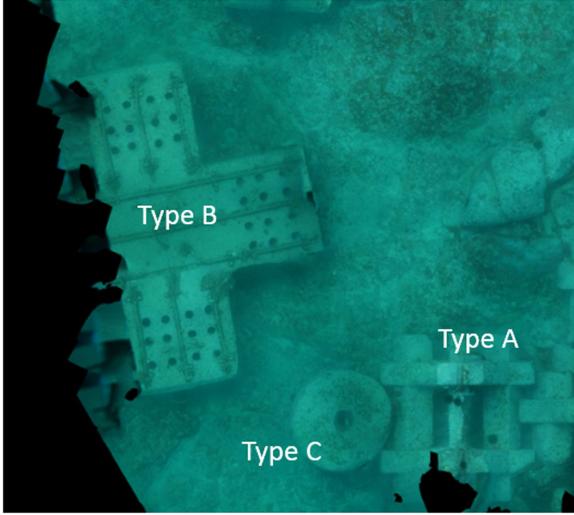


Fig. 32. Optical image of artificial reefs.

The Z coordinate of the AUV is measured by a depth meter. The pan and tilt angles of the AUV are assumed to be zero under the assumption that the horizontal attitude of the AUV is stable. The image of the AUV of our field experiment is shown in 30(c).

The AUV was equipped with DIDSON and captured successive sonar images of the seafloor at our experimental site while moving along the path, as shown in Fig. 31. The AUV captures a sonar image synchronized to its navigation sensor data update rate. For each sonar image, a set of 3-D line scan data is extracted by the proposed method, and the coordinate of the 3-D line is converted to its global coordinate equivalent in which the AUV position is defined. When the AUV repeats this line scanning for every sonar image, a 3-D point cloud of the seafloor can be generated. In this experiment, the tilt angle of DIDSON was set to 20° , and r_{\min} and r_{\max} were set to 0.8380 and 10.8380 m, respectively. The altitude of the AUV varied over time; therefore, $M_{e\min}$ and $M_{e\max}$ can be changed by altitude variation of the AUV.

On the seafloor of our experimental site, many types of artificial reefs were installed. As shown in Fig. 32, we named the artificial reef types A, B, and C. We obtained 3691 sonar image frames; several are shown in Fig. 33. In the 908th sonar image, only the seafloor is shown and no artificial reefs appear. Therefore, no highlight extension occurs in this situation as ex-

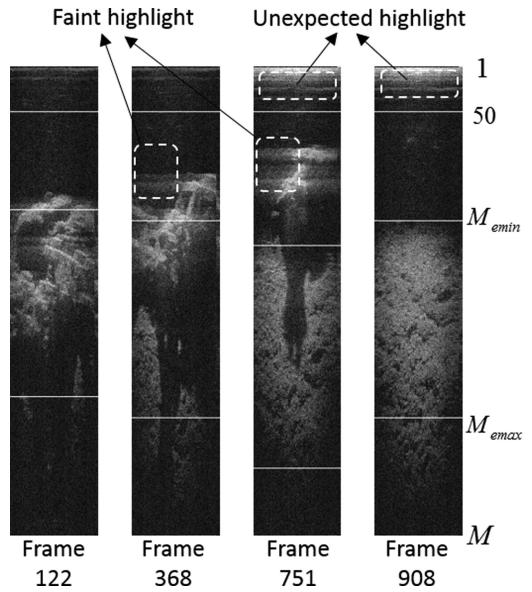


Fig. 33. Successive sonar images of field experiment.

pected. On the other hand, the 122nd, 368th, and 751st sonar images show the artificial reefs, and their highlights exceed $M_{e\min}$, which is the boundary between the effective and ineffective regions. Similar to the indoor experiment, the heights of the artificial reefs can be calculated by detecting their extended highlights.

Unlike the indoor experiment, the extended highlights of the sonar images obtained in the field experiment gradually spread from side to side, as shown in Fig. 33. This property creates faint highlights near the true extended highlights; therefore, the faint highlights may be detected as highlight extensions after applying the difference filter. However, the proposed method checks whether the detected pixel exceeds the threshold value. As a result, the proposed algorithm can effectively eliminate this type of unexpected highlight because the pixel values of the faint highlights are not high enough to be considered actual highlights. In addition, there were unexpected highlights near the top sides of the 751st and 908th sonar images, even though there was no object reflecting acoustic beams. To ignore the highlights, the pixels of the first 50 rows did not take into account the extended highlight detection, as shown in Fig. 33.

The 3-D point clouds of the artificial reefs with the trajectory of the AUV are shown in Fig. 34(a), and the heights of artificial reefs derived by the proposed method are shown in Fig. 34(b). The heights of two different points of type A were 1.137 and 1.096 m. The heights of types B and C were 0.6663 and 0.3783 m, respectively.

To verify the accuracy, we actually measured the height of the artificial reefs using a ranging pole by human divers. Fig. 35 shows video sequences of artificial reef type A taken in two different directions. According to our results, each measured heights of type A were approximately 1.215 and 1.175 m. Fig. 36 shows each video sequence of artificial reef types B and C. The heights of types B and C were measured as approximately 0.655 m and 0.36 m, respectively. Therefore, the errors of the

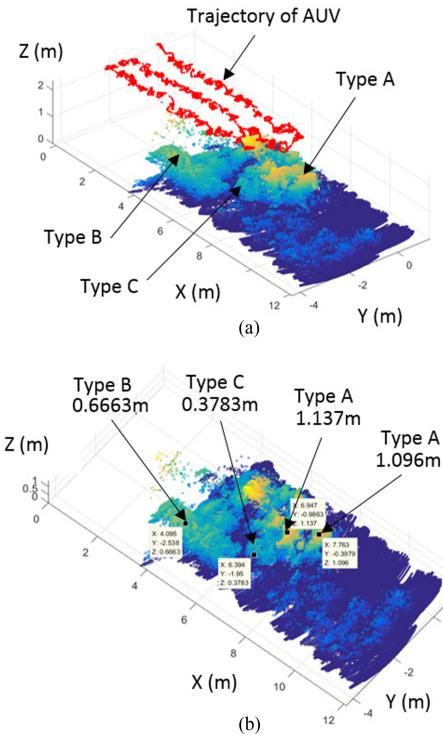


Fig. 34. Three-dimensional point clouds of the artificial reefs with the trajectory of the AUV. (a) Perspective view. (b) Height measurements.

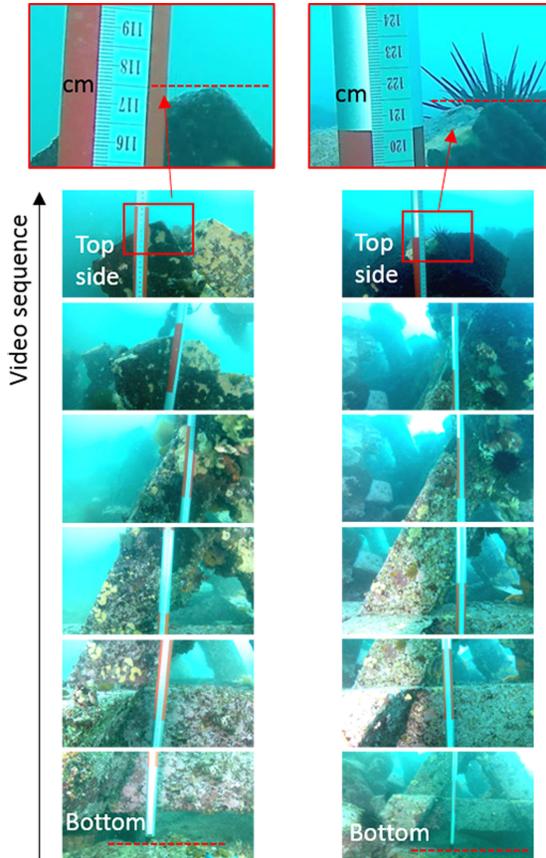


Fig. 35. Measured heights of artificial reef type A.

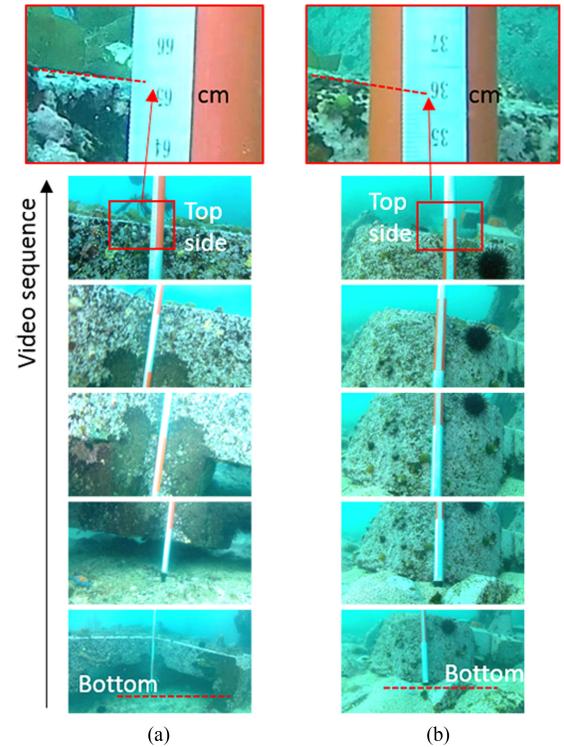


Fig. 36. Measured height of artificial reef types B and C. (a) Type B. (b) Type C.

proposed method were about a few centimeters. In the field, there exist certain errors caused by many factors such as uncalibrated sonar geometries, AUV positioning errors, and tilt angle changes because of disturbances caused by sea currents. However, the factors cannot be perfectly controlled because of the unexpected environment conditions. Even with such factors, the results show the reliability of the proposed method in the field.

In addition, one important result is that the shape of the seafloor on which the artificial reefs are placed can be recognized as 3-D point clouds. As shown in Fig. 37, the artificial reefs shown in sonar images are recognized in the 3-D point cloud map. Therefore, we claim that the proposed method is capable of generating seafloor shape information, and the 3-D data can be used for object detection in AUV applications.

Through experiments, we verified the accuracy and feasibility of the proposed method. In both a water tank and in the field, the proposed algorithm derived reasonable 3-D point cloud data. As mentioned above, the proposed algorithm requires simple image data processing, and the strategy used in the proposed method is suitable to apply to successive sonar images obtained at high frame rates. In addition, the proposed method does not require a special configuration of the ALMS for 3-D scanning, i.e., it uses normal ALMS operation to monitor the forward underwater scene. Therefore, the proposed algorithm can be used with another sonar-based algorithm at the same time. Moreover, the proposed method can be used for AUV operation because the 3-D point cloud data can provide cues for underwater object detection and obstacle avoidance.

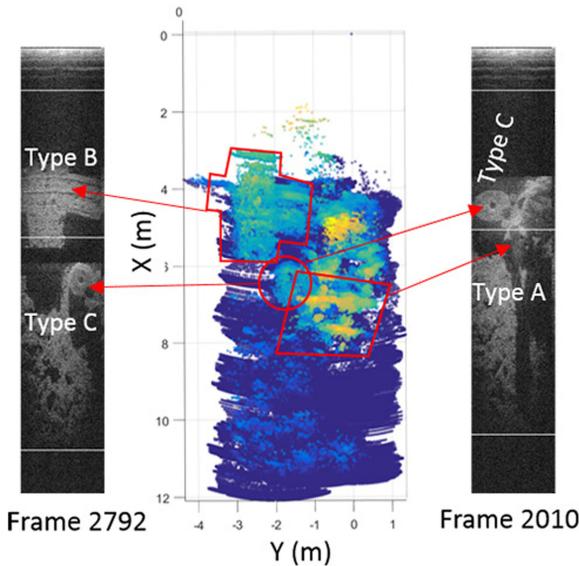


Fig. 37. Comparison between artificial reefs shown in sonar images and 3-D point clouds.

VI. CONCLUSION

In this study, we propose a method for 3-D line scanning based on ALMS. 3-D point cloud seafloor data can be generated by accumulating the 3-D line scan data of each sonar image obtained using ALMS. Unlike existing methods, the proposed method is not based on a stereo imaging technique, and therefore it can be used in unstructured environments where image matching is difficult to execute. The proposed method can be implemented using simple image processing, and therefore it is suitable for real-time AUV operations. Based on the reasonable accuracy of the 3-D point cloud data generated by the proposed algorithm, it can be used for long-range underwater object detection in AUV-based underwater investigation.

REFERENCES

- [1] T. Maki, Y. Sato, T. Matsuda, A. Kume, T. Sakamaki, and T. Ura, "AUV Tri-TON—A hover-capable platform for 3D visualization of complicated surfaces," in *Proc. IEEE Int. Symp. Underwater Technol.*, 2013, pp. 1–6.
- [2] T. Maki, A. Kume, and T. Ura, "Volumetric mapping of tubeworm colonies in Kagoshima Bay through autonomous robotic surveys," *Deep Sea Res. Part I, Oceanogr. Res. Papers*, vol. 58, no. 7, pp. 757–767, 2011.
- [3] S. B. Williams, O. R. Pizarro, and M. V. Jakuba, "Monitoring of benthic reference sites: Using an autonomous underwater vehicle," *IEEE Robot. Autom. Mag.*, vol. 19, no. 1, pp. 73–84, Mar. 2012.
- [4] E. Bovio, D. Cecchi, and F. Baralli, "Autonomous underwater vehicles for scientific and naval operations," *Annu. Rev. Control*, vol. 30, no. 2, pp. 117–130, 2006.
- [5] Y. Nishida, T. Ura, T. Sakamaki, J. Kojima, Y. Ito, and Kangsoo Kim, "Hovering type AUV "Tuna-Sand" and its surveys on Smith caldera in Izu-Ogasawara ocean area," in *Proc. OCEANS Conf.*, San Diego, CA, USA, 2013, pp. 1–5.
- [6] J. Pyo, H. Cho, H. Joe, T. Ura, and S. C. Yu, "Development of hovering type AUV "Cyclops" and its performance evaluation using image mosaicing," *Ocean Eng.*, vol. 109, pp. 517–530, 2015.
- [7] A. Bodenmann, B. Thornton, T. Ura, and M. Sangekar, "in Proc. Pixel mapping for generation of 3D coloured seafloor bathymetry using a single camera," in *Proc. IEEE/OES Auton. Underwater Vehicles*, Monterey, CA, USA, 2010, pp. 1–9.
- [8] O. Pizarro and H. Singh, "Toward large-area mosaicing for underwater scientific applications," *IEEE J. Ocean. Eng.*, vol. 28, no. 4, pp. 651–672, Oct. 2003.
- [9] N. R. Gracias and J. Santos-Victor, "Underwater video mosaics as visual navigation maps," *Comput. Vis. Image Understanding*, vol. 79, no. 1, pp. 66–91, 2000.
- [10] N. R. Gracias, S. van der Zwaan, A. Bernardino, and J. Santos-Victor, "Mosaic-based navigation for autonomous underwater vehicles," *IEEE J. Ocean. Eng.*, vol. 28, no. 4, pp. 609–624, Oct. 2003.
- [11] X. Xu and S. Negahdaripour, "Mosaic-based positioning and improved motion estimation methods for autonomous navigation," *IEEE J. Ocean. Eng.*, vol. 27, no. 1, pp. 79–99, Jan. 2002.
- [12] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging system," *IEEE J. Ocean. Eng.*, vol. 15, no. 2, pp. 101–111, Apr. 1990.
- [13] T. Ura, "Observation of deep seafloor by autonomous underwater vehicle," *Indian J. Geo-Marine Sci.*, vol. 42, no. 8, pp. 1028–1033, 2013.
- [14] M. Sangekar, B. Thornton, and T. Ura, "Wide area seafloor observation using an autonomous landing vehicle with adaptive resolution capability," in *Proc. OCEANS Conf.*, Hampton Roads, VA, USA, 2012, pp. 1–9.
- [15] G. Marani, S. K. Choi, and J. Yuh, "Underwater autonomous manipulation for intervention missions AUVs," *Ocean Eng.*, vol. 36, no. 1, pp. 15–23, Jan. 2009.
- [16] H. Cho, J. Gu, H. Joe, A. Asada, and S. C. Yu, "Acoustic beam profile-based rapid underwater object detection," *J. Mar. Sci. Tech.*, vol. 20, no. 1, pp. 180–197, Mar. 2015.
- [17] E. O. Belcher and D. C. Lynn, "Acoustic, near-video-quality images for work in turbid water," *Underwater Intervention*, Houston, TX, USA, Jan. 2000.
- [18] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2010, pp. 4396–4403.
- [19] S. Negahdaripour, "On 3-D motion estimation from feature tracks in 2-D FS sonar video," *IEEE Trans. Robot.*, vol. 29, no. 4, pp. 1016–1030, Aug. 2013.
- [20] H. Assalih, "3D reconstruction and motion estimation using forward looking sonar," Ph.D. dissertation, School Eng. Phys. Sci., Heriot-Watt Univ., Edinburgh, U.K., 2013.
- [21] M. Babaei and S. Negahdaripour, "3-D object modeling from 2-D occluding contour correspondences by opti-acoustic stereo imaging," *Comput. Vis. Image Understanding*, vol. 132, pp. 56–74, Mar. 2015.
- [22] S. C. Yu, "Development of real-time acoustic image recognition system using by autonomous marine vehicle," *Ocean Eng.*, vol. 35, pp. 90–105, Jan. 2008.
- [23] M. D. Aykin and S. Negahdaripour, "On 3-D target reconstruction from multiple 2-D forward-scan sonar views," in *Proc. OCEANS Conf.*, Genova, Italy, Sep. 2015, pp. 1–10.
- [24] N. Kumar, U. Mitra, and S. S. Narayanan, "Robust object classification in underwater sidescan sonar images by using reliability-aware fusion of shadow features," *IEEE J. Ocean. Eng.*, vol. 40, no. 3, pp. 592–606, Jul. 2015.
- [25] V. Murino and A. Trucco, "Three-dimensional image generation and processing in underwater acoustic vision," *Proc. IEEE*, vol. 88, no. 12, pp. 1903–1948, Dec. 2000.
- [26] S. Negahdaripour, "Stereo FS sonar imaging for underwater 3-D reconstruction," in *Proc. OCEANS Conf.*, Seattle, WA, USA, Sep. 2010.
- [27] N. Brahim, D. Gueriot, S. Daniel, and B. Solaiman, "3D reconstruction of underwater scenes using DIDSON acoustic sonar image sequences through evolutionary algorithms," in *Proc. OCEANS Conf.*, Santander, Spain, 2011.
- [28] A. Trucco and S. Curletto, "Extraction of 3D information from sonar image sequences," *IEEE Trans. Syst., Man, Cybern. Part B, Cybern.*, vol. 33, no. 4, pp. 687–699, Aug. 2003.
- [29] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [30] N. H. Vilarnau, "Forward-looking sonar mosaicing for underwater environments," Ph.D. dissertation, Dept. Architecture Comput. Technol, Univ. Girona, Girona, Spain, 2014.
- [31] M. Aykin and S. Negahdaripour, "On feature matching and image registration for two-dimensional forward-scan sonar imaging," *J. Field Robot.*, vol. 30, no. 4, pp. 602–623, 2013.
- [32] E. Coiras, Y. Petillot, and D. M. Lane, "Multiresolution 3-d reconstruction from side-scan sonar images," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 382–390, Feb. 2007.

- [33] K. Bikonis, M. Moszynski, and Z. Lubniewski, "Application of shape from shading technique for side scan sonar images," *Polish Maritime Res.*, vol. 30, pp. 39–44, 2013.
- [34] M. D. Aykin and S. Negahdaripour, "Forward-look 2-D sonar image formation and 3-D reconstruction," in *Proc. OCEANS Conf.*, San Diego, CA, USA, 2013, pp. 1–10.
- [35] E. O. Belcher, W. L. J. Fox, and W. H. Hanot, "Dual-frequency acoustic camera: A candidate for an obstacle avoidance, gap-filler, and identification sensor for untethered underwater vehicles," in *Proc. OCEANS Conf.*, Oct. 2002, pp. 2124–2128.
- [36] C. Xu and A. Asada, and K. Abukawa, "A method of generating 3D views of aquatic plants with DIDSON," in *Proc. IEEE Symp. Underwater Technol.*, Apr. 2011, pp. 1–5.
- [37] K. Mizuno, C. Xu, A. Asada, K. Abukawa, and M. Yamamuro, "Species classification of submerged aquatic plants using acoustic images in shallow lakes," in *Proc. IEEE Symp. Underwater Technol.*, Mar. 2013, pp. 1–5.
- [38] K. Mizuno and A. Asada, "Three dimensional mapping of aquatic plants at shallow lakes using 1.8 MHz high-resolution acoustic imaging sonar and image processing technology," in *Proc. IEEE Int. Ultrasonics Symp.*, Sep. 2014, pp. 1384–1387.
- [39] Sound Metrics, Bellevue, WA, USA. [Online]. Available: www.soundmetrics.com, Accessed on: Jan. 6, 2017.
- [40] B. K. P. Horn, "Shape from shading: A method for obtaining the shape of a smooth opaque object from one view," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1970.
- [41] S. Negahdaripour, "On 3-D scene interpretation from F-S sonar imagery," in *Proc. OCEANS Conf.*, Virginia Beach, VA, USA, Oct. 2012.
- [42] A. Bodenmann, B. Thornton, and T. Ura, "Generation of high-resolution three-dimensional reconstructions of the seafloor in color using a single camera and structured light," *J. Field Robot.*, vol. 34, no. 5, pp. 833–851, Aug. 2017.
- [43] S. Tetlow and J. Spours, "Three-dimensional measurement of underwater work sites using structured laser light," *Meas. Sci. Technol.*, vol. 10, no. 12, pp. 1162–1167, 1999.
- [44] H. Kondo, T. Maki, T. Ura, Y. Nose, T. Sakamaki, and M. Inaishi, "Relative navigation of an autonomous underwater vehicle using a light-section profiling system," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sendai, Japan, 2004, pp. 1103–1108.
- [45] G. Inglis, C. Smart, I. Vaughn, and C. Roman, "A pipeline for structured light bathymetric mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2012, pp. 4425–4432.
- [46] C. Roman, G. Inglis, and J. Rutter, "Application of structured light imaging for high resolution mapping of underwater archaeological sites," in *Proc. OCEANS Conf.*, Sydney, NSW, Australia, 2010.
- [47] A. J. Kenny, I. Cato, M. Desprez, G. Fader, R. T. E. Schüttenhelm, and J. Side, "An overview of seabed-mapping technologies in the context of marine habitat classification," *ICES J. Marine Sci.*, vol. 60, pp. 411–418, 2003, doi:10.1016/S1054-3139(03)00006-7.
- [48] M. F. Fallon, J. Folkesson, H. McClelland, and J. J. Leonard, "Relocating underwater features autonomously using sonar-based SLAM," *IEEE J. Ocean. Eng.*, vol. 38, no. 3, pp. 500–513, Jul. 2013.
- [49] N. Hurtós, X. Cufí, and J. Salvi, "A novel blending technique for two-dimensional forward-looking sonar mosaicing," *IEEE Ocean. Eng. Soc. Newslett.*, 2014, pp. 27–33.
- [50] H. Cho, J. Pyo, and S. C. Yu, "Drift error reduction based on the sonar image prediction and matching for underwater hovering," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8566–8577, Dec. 2016.
- [51] S. Negahdaripour, "Calibration of DIDSON forward-scan acoustic video camera," in *Proc. OCEANS Conf.*, Washington, DC, USA, 2005, pp. 1287–1294.



Hyeonwoo Cho (M'15) received the B.E degree from Hanyang University, Seoul, South Korea, in 2006 and the Ph.D. degree from Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2011, both in electrical engineering.

Since 2014, he has been a Research Professor in the Department of Creative IT Engineering, POSTECH. He is the CEO of Pohang Ocean System Lab. Inc. His research interests include underwater object detection and recognition algorithms.

Dr. Hyeonwoo is the Secretary of Ocean Engineering Society Korea Chapter.



Byeongjin Kim (M'15) received the B.E degree in electrical engineering from Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2015, where he is currently working toward the Ph.D. degree in robotics.

He is a member of Hazardous and Extreme Environment Robotics Laboratory, POSTECH. His research interest include underwater robotics and underwater sonar image processing for navigation.



Son-Cheol Yu (M'05) received B.E. degree in naval architecture and ocean engineering from Inha University, Incheon, South Korea, in 1998 and the M.E. and Ph.D. degrees in ocean and environmental engineering from the University of Tokyo, Tokyo, Japan, in 2000 and 2003, respectively.

He was a Researcher in the Department of Mechanical Engineering, University of Hawaii, Honolulu HI, USA, from 2004 to 2007 and a faculty member in the Department of Mechanical Engineering, Pusan National University, Busan, South Korea, from 2008 to 2009. Since 2010, he has been a faculty member in the Department of Creative IT Engineering, Pohang University of Science and Technology (POSTECH), Pohang, South Korea. He was a Guest Investigator at the Woods Hole Oceanographic Institution, Woods Hole, MA, USA. His current research focuses on the autonomy of underwater vehicles and intelligent sensing.

Prof. Yu was the founder and Chair of the Oceanic Engineering Society Korea Chapter.